Pis'ma v ZhETF, vol.66, iss.3, pp.200 - 205

SEMICLASSICAL THEORY OF COULOMB ANOMALY

S.Levitov[1], *A.V.Shytov*[2]

Massachusetts Institute of Technology,
Cambridge, MA-02139, USA

L.D.Landau Institute for Theoretical Physics RAS
117334 Moscow, Russia

Submitted 12 May 1997
Resubmitted 15 July 1997

An effective action approach for the problem of Coulomb blocking of tunneling is discussed. The method is applied to the strong coupling problem arising near zero bias, where perturbation theory diverges. We find an instanton for the electrodynamics in imaginary time, and express the anomaly through exact conductivity of the system $\sigma(\omega, q)$ and exact interaction.

PACS: 72.10.Bg, 73.50.-h

Suppression of tunneling conductivity near zero bias (called "zero-bias anomaly") is known to be a signature of interaction in the system. The anomaly has been studied in metals and semiconductors since the early sixties [1]. Initially, the anomaly was attributed to the Kondo effect, but it was realized later that a much more common mechanism is Coulomb blocking of tunneling. A perturbation theory of this effect was developed by Altshuler, Aronov and Lee [2]. They show that in a diffusive conductor tunneling blocking increases at small bias, leading to a singularity in tunneling conductivity. The theory has been thoroughly tested experimentally [3].

In the last years the interest shifted to the systems with strong Coulomb effects, such as disordered metals and semiconductors near metal-insulator transition [4]. It is found that the Coulomb anomaly is sharply enhanced near the transition, providing a test of electron correlations. Another important development is due to studies by Ashoori et al. of electron tunneling into a two-dimensional metal in magnetic field [5]. In this experiment it is found that at certain magnetic field the zero-bias anomaly abruptly increases and transforms to a "soft Coulomb gap". It has been pointed out [6] that this transition is induced by disorder. Eisenstein et al. studied systems with higher mobility [6, 7], where tunneling is almost entirely blocked below certain threshold bias. These findings caused a lot of theoretical work [7] aimed at relating the tunneling anomaly with microscopic models. Also, the anomaly has been shown to be particularly interesting in the Quantum Hall system [8].

Below we discuss an effective action theory [9] that treats the anomaly as cooperative tunneling. We obtain tunneling exponent in terms of the action of spreading charge density and express it through the actual conductivity $\sigma(\omega, q)$ of the system. The treatment is non-perturbative and remains accurate both in the weak and strong coupling regimes.

Qualitative discussion. Tunneling of an electron into a metal involves two steps: traversing the barrier, followed by spreading within the metal. Typically, the traversal time is very short compared to the spreading time. Accordingly, it

[1] e-mail: levitov@mit.edu
[2] e-mail: shytov@mit.edu

is legitimate to separate the tunneling into single-electron and many-electron steps, and to treat them separately. The first contribution is simply the transmission coefficient of the barrier, constant at small bias. The second contribution is more interesting, since it is a collective effect involving motion of a large number of electrons in order to accomodate the new electron. At low bias the latter effect completely controls the tunneling rate.

Let us illustrate the effect of charge relaxation on the tunneling rate by using the example of a two-dimensional conducting plane [10]. Charge relaxation in a two-dimensional conductor is a classic electrodynamics problem studied by Maxwell who gave a solution in terms of a moving image charge [11]. In this problem, a point charge e is injected into a conducting sheet with conductivity σ, and one is interested in the time-dependence of the density and potential of spreading charge. Maxwell's solution is that the potential within the sheet is given by that of a point charge e moving along the normal to the plane with the velocity $v = 2\pi\sigma$. The size of the charge cloud grows as $r(t) \sim vt$. Let us consider the Coulomb part of the action for the charge:

$$\mathcal{S}(t) \sim \int_{t_0}^{t} \frac{e^2}{r(t')} dt' = \frac{e^2}{2\pi\sigma} \ln\left(\frac{t}{t_0}\right). \qquad (1)$$

Semiclassically, the action (1) must be added to the under-barrier action. The divergence of (1) at $t \to \infty$ indicates that at large times the action (1) dominates the tunneling rate. We will see that the time of spreading diverges at small bias, $t_* = e/\sigma V$. From that, near zero bias the tunneling acquires a suppression factor of $\exp\left(-\frac{1}{\hbar}\mathcal{S}(t_*)\right)$. The estimate (1) showing that the action diverges at small bias means that the semiclassical treatment is meaningful even for a well conducting metal. However, in the diffusive limit the estimate (1) does not agree with the perturbation theory. We shall see that the reason is that the main part of the action is rather Ohmic than Coulomb, and that after writing the action properly the semiclassical method completely recovers the pertubation theory result.

The relevance of the semiclassical picture can be justified by a more general argument, not involving specific features of 2D relaxation. Let us consider effects caused by one electron tunnelling into a metal. Since the barrier traversal time is much shorter than the relaxation time in the metal, while the electron traverses the barrier other electrons practically do not move. Thus instantly a large electrostatic potential is formed, both due to the tunneling electron itself, and due to the screening hole left behind. At small bias eV, the increase of electrostatic energy by an amount much larger than the bias indicates that right after the electron transfer the system is found in a classically forbidden state "under the Coulomb barrier." In order to accomplish tunneling, the charge yet has to spread over a large area, so that the potential of the charge fluctuation is reduced below eV. If the conductivity becomes smaller, charge spreading takes longer time, and thus the action of the cooperative under-barrier motion becomes much greater than $\hbar$.

The form of the action. For the electrodynamics problem the action can be written in terms of charge and current densities $\rho(\mathrm{r}, t)$ and $\mathrm{j}(\mathrm{r}, t)$. Full action would also contain electromagnetic potentials, but in the quasistationary electrodynamics, $c \to \infty$, which we always assume below, the potentials are "slaved" to charges, and thus can be integrated out. Also, since the contribution to the spreading charge action mainly arises from large space and time scale, we assume that local deviation from equilibrium is small. Therefore, one can expand the action in powers of $\rho(\mathrm{r}, t)$ and $\mathrm{j}(\mathrm{r}, t)$, and keep only quadratic terms. Naturally, the

quadratic action must reproduce correctly the classical electrodynamics equations: the Ohm's law and charge continuity.

In fact, this requirement is entirely sufficient to determine the form of the action. However, it is more tutorial to argue in the following way. We are going to use the action to study the dynamics in imaginary time. Therefore, the action is precisely the one that appears in the quantum partition function. The latter action expanded up to quadratic terms in charge and current density must yield correct Nyquist spectrum of equilibrium current fluctuations:

$$\langle\langle \mathrm{g}^{\alpha}_{\omega,q}\mathrm{g}^{\beta}_{-\omega,-q}\rangle\rangle = \sigma_{\alpha\beta}|\omega| + \sigma_{\alpha\alpha'}D_{\beta\beta'}q_{\alpha'}q_{\beta'} \ . \tag{2}$$

Here $\mathrm{g} = \mathrm{j} + \hat{D}\nabla\rho$ is external current and $D_{\alpha\beta}$ is the tensor of diffusion constants related to the conductivity tensor by the Einstein's formula: $\hat{\sigma} = e^2\nu\hat{D}$, where $\nu = dn/d\mu$ is compressibility. Generally, both $\hat{\sigma}$ and $\hat{D}$ are functions of the frequency and momentum. For simplicity, below we assume zero temperature and isotropic conductivity.

The requirement of matching equilibrium current fluctuations is essentially equivalent to the fluctuation-dissipation theorem. Thus, the form of the action is fixed by response functions of the system. In imaginary time it reads:

$$\mathcal{S} = \frac{1}{2}\int\int d^4x_1 d^4x_2 \left[\mathrm{g}_1^T \hat{K}_{x_1-x_2}\mathrm{g}_2 + \frac{\delta_{12}\rho_1\rho_2}{|\mathrm{r}_1-\mathrm{r}_2|}\right] \tag{3}$$

where $x_{1,2} = (t_{1,2}, \mathrm{r}_{1,2})$, and $\delta_{12} = \delta(t_1 - t_2)$. The kernel $\hat{K}_{t,\mathrm{r}}$ is related to the current-current correlator,

$$(K^{-1}_{\omega,q})_{\alpha\beta} = \langle\langle \mathrm{g}^{\alpha}_{i\omega,q}\mathrm{g}^{\beta}_{-i\omega,-q}\rangle\rangle \tag{4}$$

given by (2), where $\hat{\sigma}$ and $\hat{D}$ are functions of the Matsubara frequency obtained from the real frequency functions by the usual analytic continuation. We take Coulomb interaction in the second term of the action (3) as non-retarded because we are going to study systems with relatively low conductivity, and thus slow charge relaxation.

Instanton. To evaluate the tunnelling rate, we construct an instanton, i.e., a "bounce path" in imaginary time[12, 13]. Among bounce paths symmetric in time, $\rho(\mathrm{r},t) = \rho(\mathrm{r},-t)$, $\mathrm{j}(\mathrm{r},t) = -\mathrm{j}(\mathrm{r},-t)$, we shall find the least action path which will give a semiclassical tunneling rate exponent.

In order to get equations of motion from the variational principle one notes that the action (3) contains the charge and current densities as *independent* variables, since Eq.(3) was derived by matching with the equilibrium fluctuations in the *grand canonical ensemble* where charge is not conserved. Therefore, one must supply the action (3) with the charge continuity constraint: $\dot{\rho} + \nabla\cdot\mathrm{j} = \mathcal{J}(\mathrm{r},t)$, where

$$\mathcal{J}(\mathrm{r},t) = e\delta(\mathrm{r})(\delta(t+\tau) - \delta(t-\tau)) \ . \tag{5}$$

The charge source $\mathcal{J}(\mathrm{r},t)$ describes electron injected at $\tau = 0$, $t = -\tau$, and taken back at $t = \tau$, at the same point. It is assumed that tunneling occurs at one point (impurity or a "hot spot").

Charge continuity is incorporated in the action by a Lagrange mutiplier:

$$\mathcal{S}_{total} = \mathcal{S}(\rho,\mathrm{j}) + \phi(\mathrm{r},t)\left(\dot{\rho} + \nabla\mathrm{j} - \mathcal{J}(\mathrm{r},t)\right) \ , \tag{6}$$

where $\phi(\mathrm{r},t)$ is an independent variable of the problem. For the least action path, the variation of $\mathcal{S}_{total}$ relative to infinitesimal change $\delta\phi$, $\delta\rho$, and $\delta\mathrm{j}$ vanishes. (Note that $\delta\dot{\rho} + \nabla\cdot\delta\mathrm{j} = 0$.) After eliminating ϕ we get standard electrodynamics equations:

$$\text{(i)} \qquad \dot{\rho} + \nabla\cdot\mathrm{j} = \mathcal{J}(\mathrm{r},t) \ ; \qquad \text{(ii)} \qquad \mathrm{j} + D\nabla\rho = \hat{\sigma}(\omega,\mathrm{q})\,\mathrm{E} \ ;$$

(iii) $$\mathbf{E}(\mathbf{r},t) = -\nabla_{\mathbf{r}} \int dr' \rho(r',t) U(|\mathbf{r}-\mathbf{r}'|) \ . \qquad (7)$$

Eqs. (7) describe evolution in imaginary time, i.e., charge spreading under the Coulomb barrier arising from selfinteraction.

Then one must solve Eq.(7) for ρ and j, and to compute the action (3). For a spatially homogeneous system, by using Fourier transform, we get

$$\rho(\omega,\mathbf{q}) = \frac{\mathcal{J}(\omega)}{|\omega| + D\mathbf{q}^2 + \sigma \mathbf{q}^2 U_q} \quad , \quad \mathbf{j}(\omega,\mathbf{q}) = -i\hat{K}^{-1}(\omega,\mathbf{q})\,\mathbf{q}\,U_q\,\rho(\omega,\mathbf{q}) \ , \qquad (8)$$

where $U_{\mathbf{q}}$ is the Coulomb potential formfactor. Then the action (3) becomes

$$\mathcal{S}_0(\tau) = \frac{1}{2}\sum_{\omega,\mathbf{q}} \frac{|\mathcal{J}(\omega)|^2}{|\omega| + Dq^2} \frac{U_q}{|\omega| + Dq^2 + \sigma q^2 U_q} \quad , \qquad (9)$$

depending on τ through Fourier component of the charge source: $\mathcal{J}(\omega) = 2ie \sin \omega\tau$. **Accomodation time:** Finally, to obtain total action of the system we subtract the term $2eV\tau$ from the spreading charge action $\mathcal{S}_0(\tau)$ in order to account for the voltage source work: $\mathcal{S}(\tau) = \mathcal{S}_0(\tau) - 2eV\tau$. By that the energy conservation is assured. Then one optimizes $\mathcal{S}(\tau)$ in τ:

$$\partial \mathcal{S}_0(\tau)/\partial\tau|_{\tau=\tau_*(V)} = 2eV \qquad (10)$$

The optimal time τ_* can be interpreted as the charge accomodation time. By plugging τ_* found from Eq.(10) into the tunneling rate exponent one gets tunneling conductivity $G(V) = G_0 e^{-\mathcal{S}(\tau_*(V))/\hbar}$. The accuracy of this calculation relies on the assumption that the time τ_* is larger than the barrier traversal time τ_f. This is true whenever there is an anomaly: if τ_* were $\approx \tau_f$, then $\mathcal{S}(\tau_*) \approx \hbar$ and there would be no tunneling suppression.

A more rigorous way to proceed is to write the equal-point Green's function $\tilde{G}(t) = G_0(t)e^{-S_0(t/2)}$, where $G_0(t) = \nu_0/(t+i0)$ is the equal-point free electron Green's function. The tunneling current can be then written as

$$I = e\mathrm{T}\,\mathrm{Im} \int \tilde{G}_1(t)\tilde{G}_2(-t)e^{ieVt}dt \ , \qquad (11)$$

where T is barrier transmission constant, and $G_{1,2}$ are Green's functions in the leads. The equivalence of Eq.(11) and the above procedure (10) can be established by evaluating the integral (11) by the steepest decent method.

Using (9) and (11) one can study tunneling anomaly in different systems. In doing so, while calculating $\tau_*(V)$ and $\mathcal{S}(\tau_*)$ it is essential to check the selfconsistency of the assumption that $\tau_* \gg \tau_f$. For example, this assumption will not be fulfilled in a *clean* metal, i.e., in Fermi liquid without disorder ($D > 1$). The reason is that the conductivity of an ideal conductor is $\sigma(\omega) = ine^2/m\omega$. In this case Eq.(9) gives $\mathcal{S}_0(\tau) \approx \hbar$ at any $\tau \gg \tau_f$, and henceforth $\tau_* \simeq \tau_f$. This indicates absence of the anomaly in a clean metal, the result familiar from the Fermi liquid picture. On the contrary, for a one-dimensional metal $\mathcal{S}_0(\tau) \sim \ln \tau/\tau_f$, which leads to the power-law anomaly known from the Luttinger liquid theory[14].

Comparison with previous work. For a two-dimensional metal with elastic scattering time τ_0 and non-screened Coulomb interaction we set $U_q = 2\pi/|q|$ and $\sigma(\omega,q)$ constant at $|\omega|, v_F|q| \le 1/\tau_0$. Then Eq.(9) gives

$$\mathcal{S}(\tau) = \frac{e^2}{8\pi^2\sigma} \ln\left(\frac{\tau}{\tau_0}\right) \ln\left(\tau\tau_0\sigma^2(\nu e^2)^2\right) \ . \qquad (12)$$

From Eq.(10), $\tau_* = (e/4\pi^2 V\sigma)\ln(\hbar\sigma\nu e/V)$. The theory is selfconsistent at long times, $\tau_* \geq \tau_0$, i.e., at $eV \leq e^2/\sigma\tau_0$. Then the least action is

$$\mathcal{S}(V) = \frac{e^2}{8\pi^2\sigma}\ln\left(\frac{e}{4\pi^2\sigma V\tau_0}\right)\ln\left(\frac{e\tau_0\sigma(\nu e^2)^2}{4\pi^2 V}\right) \qquad (13)$$

It is interesting to compare this result with the identical double-log dependence derived by Altshuler, Aronov, and Lee[2] for the correction to the tunneling density of states $\delta\nu(\epsilon)$. The calculation[2] assumes that $\delta\nu$ is small, $|\delta\nu| \ll \nu_0$, which is the case only for a weak disorder. It was found that $\delta\nu(\epsilon) = -\hbar^{-1}\nu_0\mathcal{S}(eV = \epsilon)$, where $\mathcal{S}(V)$ is given by (13). The main difference is that the double-log (13) has to be exponentiated to get the tunnelling density of states, while in [2] the double-log appears as a correction to the density of states. Within perturbation theory domain the two results agree.

Also, let us mention a relation with the work by Halperin, He, and Platzman[8] dealing with the anomaly in the $\nu = 1/2$ Quantum Hall state. The effect considered in this work is a shakeup of low–frequency modes of the system due to tunneling electron, assuming that the response function is described by the Chern-Simons Fermi liquid theory[15]. The anomaly was found to have the form:

$$G(V) \sim \exp(-V_0/V) \quad , \quad V_0 = 4\pi\frac{e}{\epsilon}\sqrt{\pi n} \quad , \qquad (14)$$

where $V \ll V_0$, and n is density. It is interesting to see how this result can be derived from the effective action. It has been shown[15] that conductivity of the $\nu = 1/2$ state has strong spatial dispersion: $\sigma(k) = A|k|$, $A = e^2/16\pi\epsilon\sqrt{\pi n}$. If this form is inserted in the action (9), one gets $S(\tau) = \pi\sqrt{2\tau/A}$, which leads to the tunneling rate (14).

Source–drain correlation. For tunneling between source and drain, there are separate contributions to the action due to the relaxation of the electron and hole charges on both sides of the barrier. If the barrier is thin, the positive and negative charges partially screen the field of each other, which makes their spreading correlated. In this case the least action is smaller than the sum of independent contributions of the electrodes, and thus the anomaly is weakened. For D = 2, quantitatively, the effect will be that the voltage dependence of the second log in Eq.(13) saturates at $eV \simeq V_0 = \min[\sigma_1, \sigma_2]\hbar/a$, where a is the distance between the electrodes.

This *excitonic* correlation effect can be treated straightforwardly by writing the action (3) for each electrode separately, together with the term describing interaction across the barrier. Let us consider an example of two parallel planes with different conductivities and diffusion constants, σ_1, σ_2, D_1, and D_2. It is straightforward to generalize the above procedure and to find the instanton. The least action is

$$\mathcal{S}_0(\tau) = \frac{1}{2}\sum_{\omega,q}|J(\omega)|^2\begin{bmatrix}1\\-1\end{bmatrix}^{\mathrm{T}}\left(|\omega| + \mathbf{D}q^2\right)^{-1}\mathbf{U}\left(|\omega| + \mathbf{D}q^2 + \mathbf{\Sigma}q^2\mathbf{U}\right)^{-1}\begin{bmatrix}1\\-1\end{bmatrix} \quad , \qquad (15)$$

where we use matrix notation:

$$\mathbf{D} = \begin{pmatrix}\hat{D}_1 & 0\\ 0 & \hat{D}_2\end{pmatrix} \quad , \quad \mathbf{\Sigma} = \begin{pmatrix}\hat{\sigma}_1 & 0\\ 0 & \hat{\sigma}_2\end{pmatrix} \quad , \quad \mathbf{U} = \begin{pmatrix}U_q & V_q\\ V_q & U_q\end{pmatrix} , \qquad (16)$$

$$U_q = \int e^{iqr}\frac{d^2r}{|r|} \quad , \quad V_q = \int \frac{e^{iqr}}{\sqrt{r^2+a^2}}d^2r \quad .$$

The rows and columns of the matrices (16) correspond to the two planes.

To make evaluation of $\mathcal{S}_0(\tau)$ simpler, let us assume $e^2\nu_{1,2}a \gg 1$, the condition satisfied in almost all experiments. By carrying out matrix inversion and integration

we get

$$\mathcal{S}_0(\tau) = \alpha \ln\left(\frac{\tau}{\tau_0}\right) \qquad \text{at} \qquad \tau \gg \hbar/eV_0 \ , \tag{17}$$

where

$$\alpha = \frac{e^2}{4\pi^2}\left[\frac{1}{\sigma_1}\ln\frac{4\pi\sigma a}{D_2} + \frac{1}{\sigma_2}\ln\frac{4\pi\sigma a}{D_1}\right] \ , \tag{18}$$

where $\sigma = \sigma_1\sigma_2/(\sigma_1+\sigma_2)$. If the two planes are identical,

$$\alpha = \frac{e^2}{2\pi^2\sigma}\ln 2\pi e^2\nu a \ . \tag{19}$$

Thus at $V < V_0$ the $I-V$ curve becomes the power law $I \sim V^{\alpha+1}$. The tunneling suppression in this case is weaker than for the non-screened interaction.

Conclusion. The essential feature of the above approach is that it relates the tunneling anomaly with the actual conductivity of the system. This would allow a comparison with experiment in the situations where there is no accepted model for conductivity. For example, the tunneling current as function of voltage can be taken from experiment, and directly used to find $\mathcal{S}(\tau)$ by Legendre transform. Then $\mathcal{S}(\tau)$ can be analyzed by using (9) and (11) to extract the conductivity frequency and wavevector dependence.

To summarize, we argued that the theory of the Coulomb anomaly in the regime of strong suppression of tunneling is semiclassical. The underlying reason is that the transfer of one electron across the barrier is controlled by cooperative motion of many other electrons. We treat this motion as classical electrodynamics in imaginary time, for which we construct an effective action and find an instanton path. The instanton action determines tunneling exponent as function of the bias.

---

1. B.L.Altshuler and A.G.Aronov, in: *Electron-Electron Interaction in Disordered Systems*, Eds. A.L.Efros and M.Pollak (North-Holland, 1985); R.C.Dynes and P.A.Lee, Science **223**, 355 (1984);
2. B.L.Altshuler, A.G.Aronov, and P.A.Lee, Phys. Rev. Lett., **44**, 1288 (1980).
3. R.C.Dynes and J.P.Garno, Phys. Rev. Lett. **46**, 137 (1981); Y.Imry and Z.Ovadyahu, Phys. Rev. Lett. **49**, 841 (1982).
4. J.M.Valles, R.C.Dynes, and J.P.Garno, Phys. Rev. B **40**, 7590 (1989); Phys. Rev. B **40**,6680 (1989).
5. R.C.Ashoori, J.A.Lebens, N.P.Bigelow, and R.H.Silsbee, Phys. Rev. Lett. **64**, 681 (1990); Phys. Rev. B **48**, 4616 (1993);
6. J.P.Eisenstein, L.N.Pfeiffer, and K.N.West, Phys. Rev. Lett., **69**, 3804 (1992).
7. S.R.E.Yang and A.H.MacDonald, Phys. Rev. Lett. **70**, 4110 (1993); Y.Hatsugai, P.A.Bares, and X.G.Wen, Phys. Rev. Lett. **71**, 424 (1993); P.Johansson and J.M.Kinaret, Phys. Rev. Lett. **71**, 1435 (1993); A.L.Efros and F.G.Pikus, Phys. Rev. **B 48**, 14694 (1993); C.M.Varma, A.I.Larkin, and E.Abrahams, Phys. Rev.**B 49**, 13999 (1994); S.R.Renn and B.W.Roberts, Phys. Rev.**B 50**, 7626 (1994); I.L.Aleiner, H.U.Baranger, and L.I.Glazman, Phys. Rev. Lett. **74**, 3435 (1995); Phys. Rev. **B 50**, 15465 (1994).
8. S.He, P.M.Platzman, B.I.Halperin, Phys. Rev. Lett. **71**, 777 (1993).
9. L.S.Levitov and A.V.Shytov, in: *Correlated Fermions and Transport in Mesoscopic Systems*, Coulomb Blocking of Tunneling: From Zero-bias Anomaly to Coulomb Gap, Les Arcs, Savoie, (Editions Frontieres, 1996).
10. B.Z.Spivak, unpublished (1990).
11. J.C.Maxwell, *A Treatise on Electricity and Magnetism*, vol. II, Ch. XII, Oxford: Clarendon Press, p.296, 1892.
12. J.S.Langer, Ann. Phys. **41**, 108 (1967).
13. S.Coleman, Phys. Rev. D **15**, 2929 (1977).
14. G.D.Mahan, *Many-particle physics*, Sec. 4. 4, Plenum Press, 1981.
15. B.I.Halperin, P.A.Lee, and N.Read, Phys. Rev. B **47**, 7312 (1993).